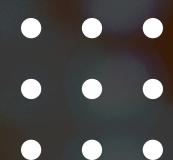




# FOREX FORECASTING WITH SENTIMENT ANALYSIS

Model Evaluation & Insights



Martí Girón, Roger Guiñón and Antonio Stanciu



• • •  
• • •  
• • •

# Problem Statement & Motivation

- High-liquidity yet volatile Forex market: predicting daily EUR/USD direction is challenging.
- Hypothesis: Combining historical price data with GDELT-derived news sentiment improves classification of SELL, HOLD, BUY signals.
- Goal: Deliver a reproducible DL pipeline to test if sentiment adds predictive value over a pure price baseline.



# Dataset & Pre-Processing

01

**Price Data:** EUR/USD OHLC from 2013-2025, transformed to log-returns.

02

**News Sentiment:** GDELT Global Event DB 22 k headlines → FinBERT polarity → daily aggregate.

03

**Merged dataset aligned on date;** NaNs forward-filled & features standardised.

04

**Sliding window of 30 days to form model inputs (return & sentiment vectors).**

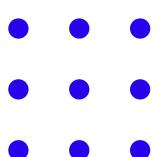


# Model 1 - XGBoost Baseline



**Tabular window flattened →  
Gradient-boosted trees (depth = 8,  
 $n = 150$ )**

**Handles non-linear interactions  
without explicit sequential  
modelling**





# Model 2 - LSTM

**3-layer PyTorch LSTM (hidden 128, dropout 0.3) fed with 30-day sequences**

**Adam optimiser, LR 5e-4, early stopping (patience = 15)**

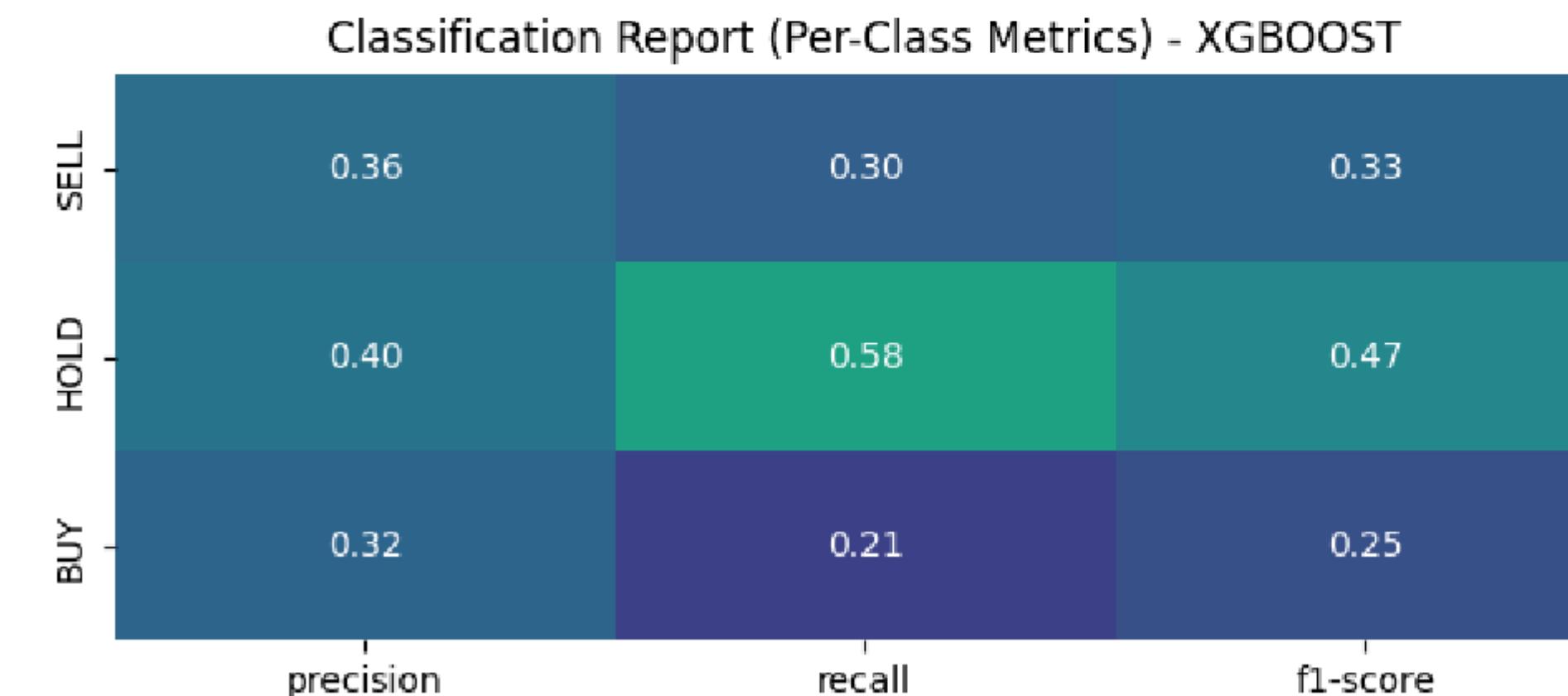
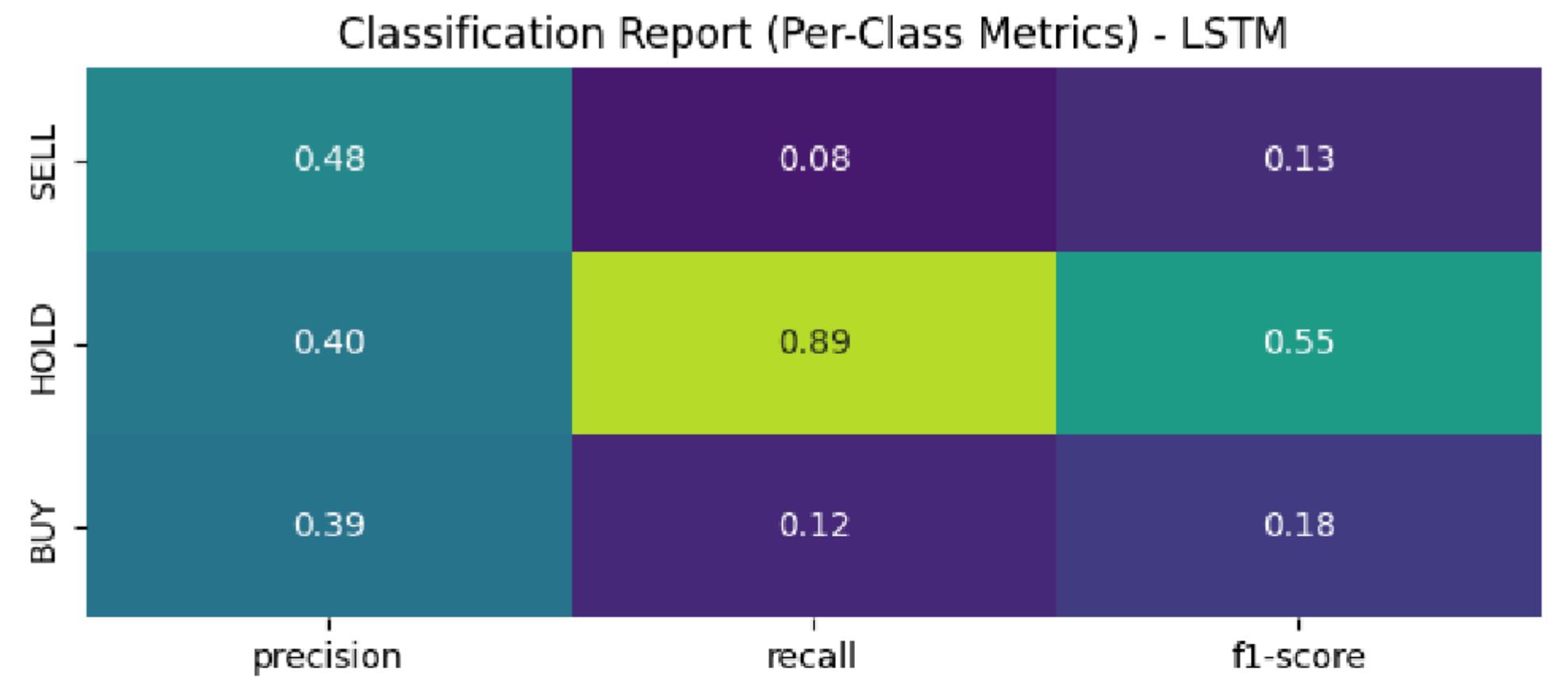
**Outputs 3-way softmax; class weighting applied to mitigate imbalance**





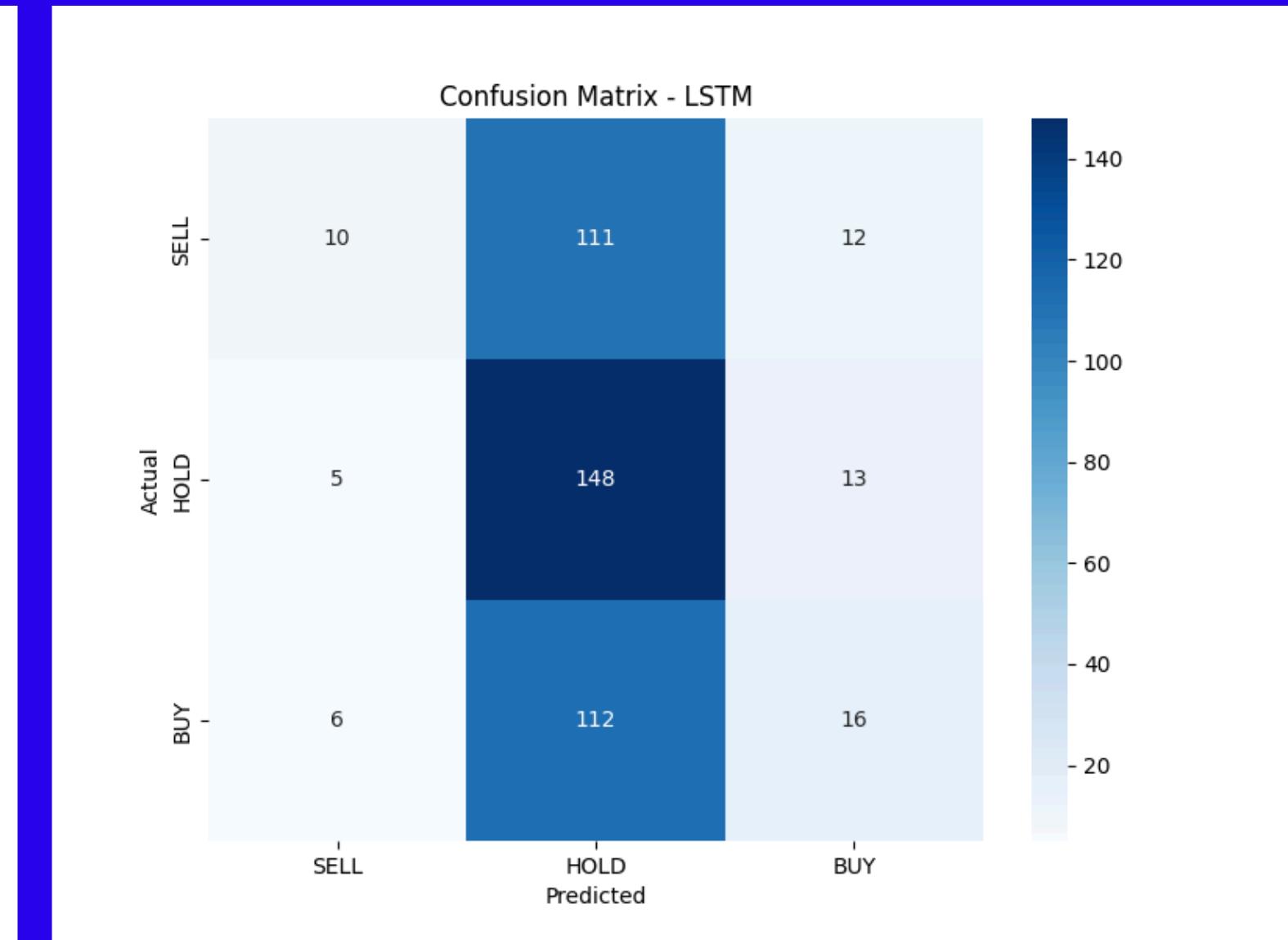
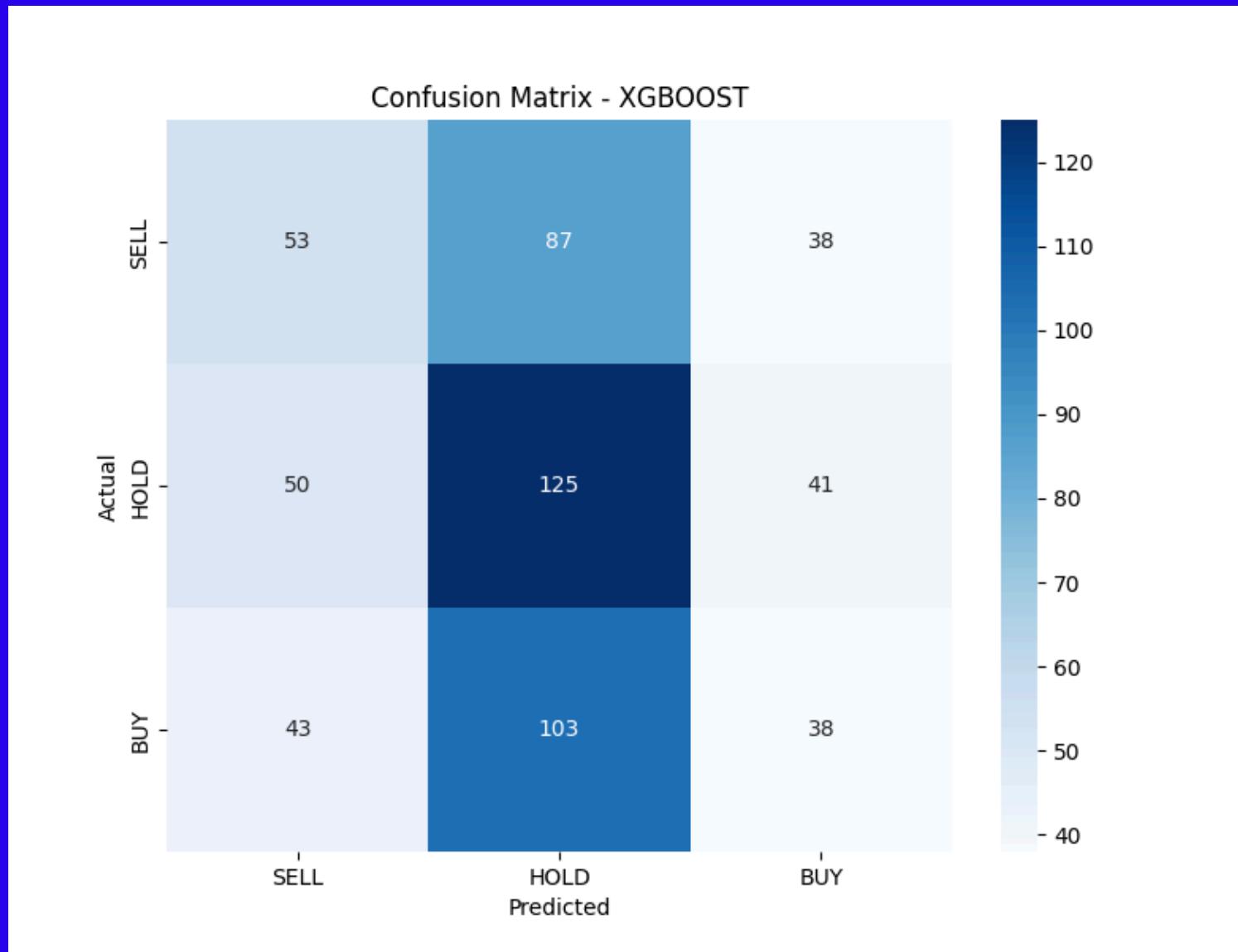
# Evaluation Metrics

• • •  
• • •  
• • •



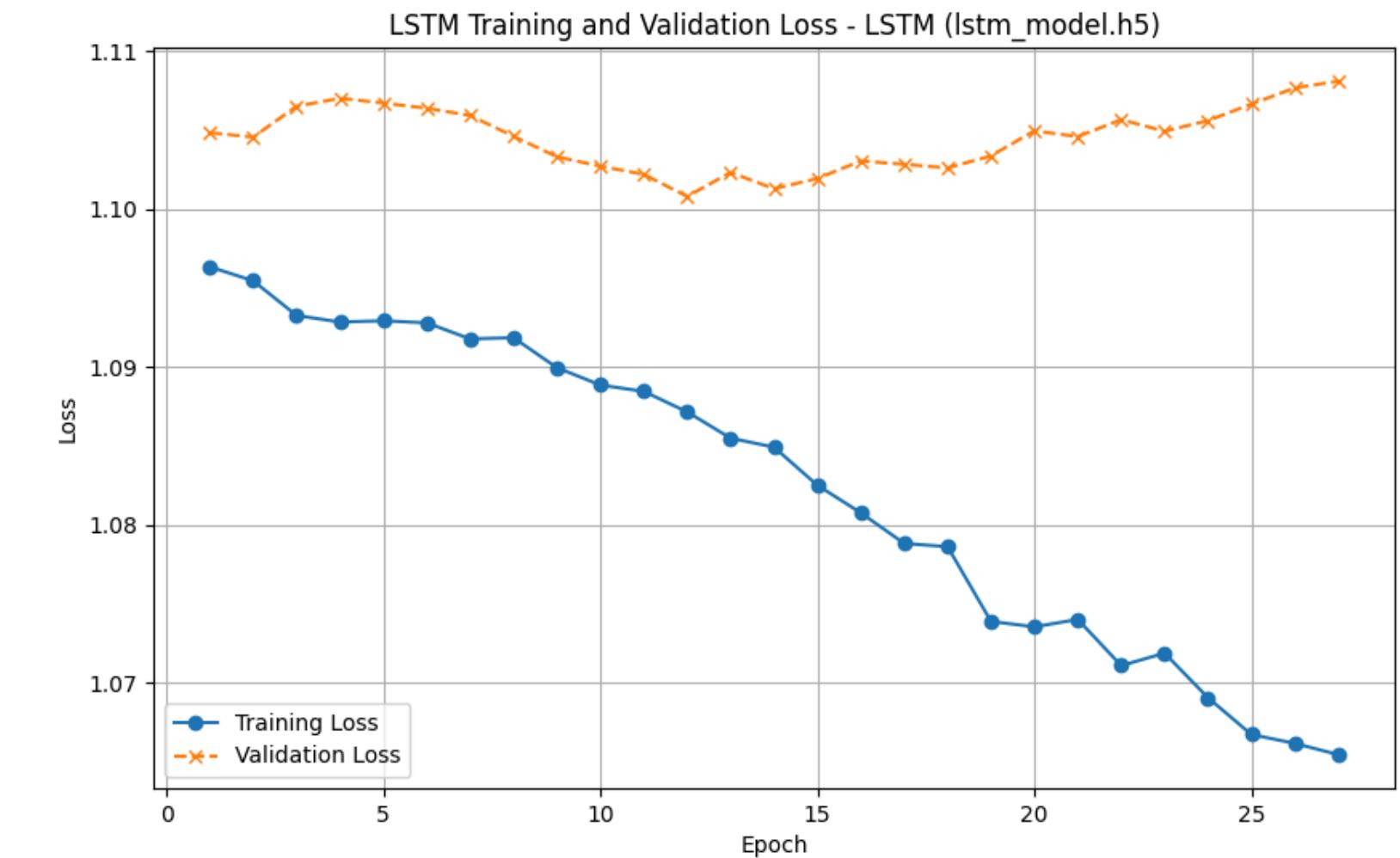
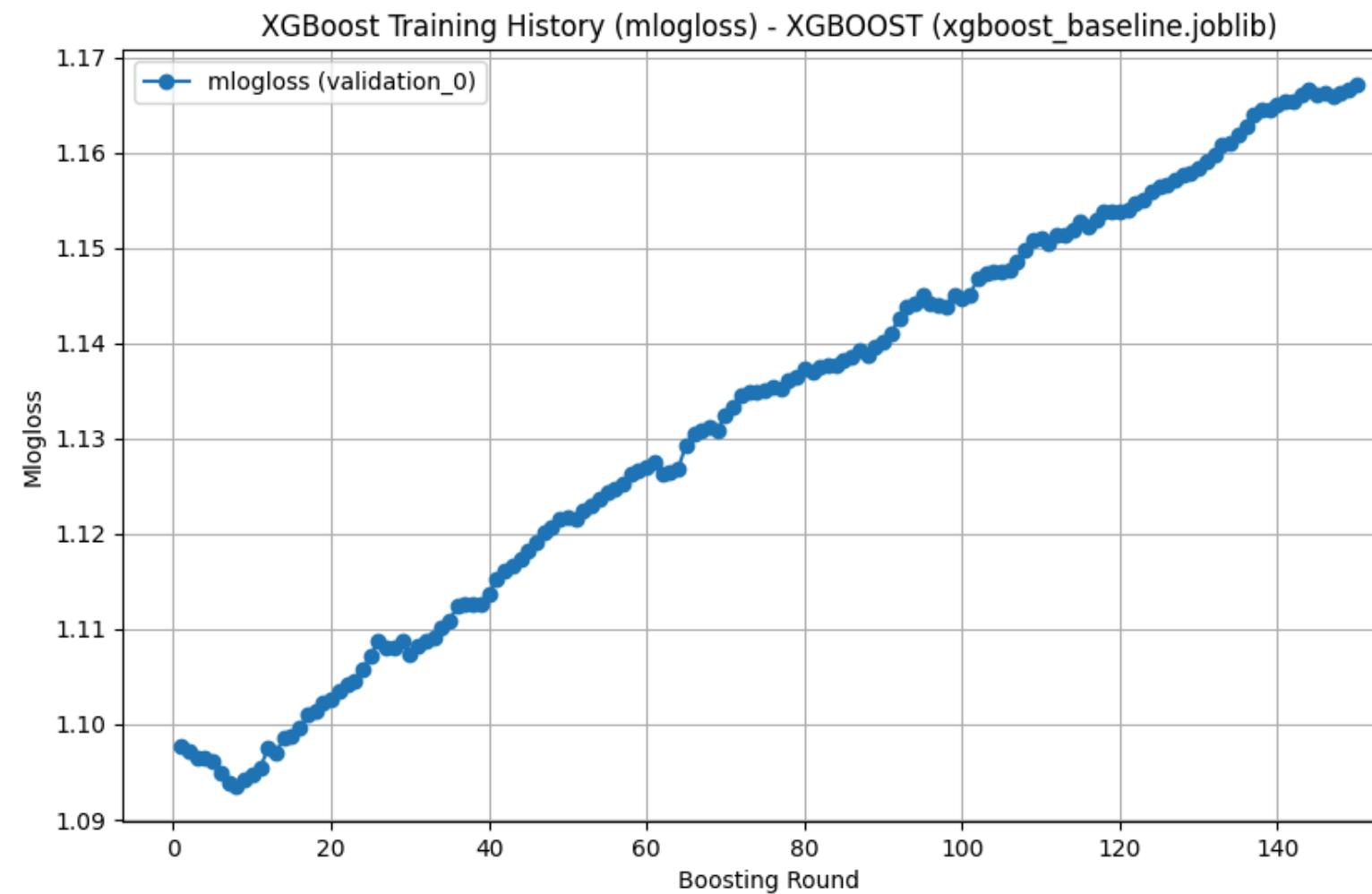


# Confusion Matrices



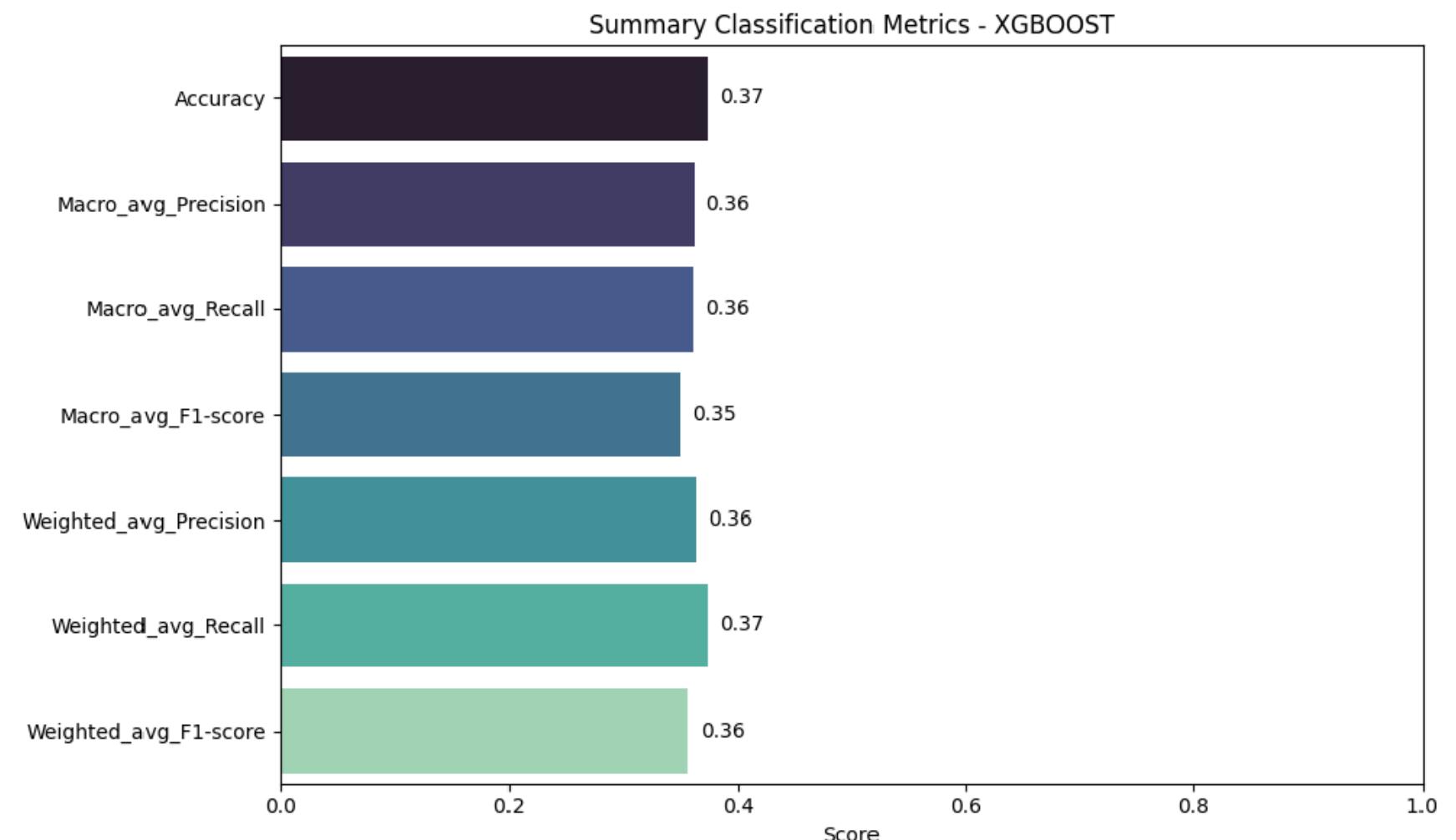
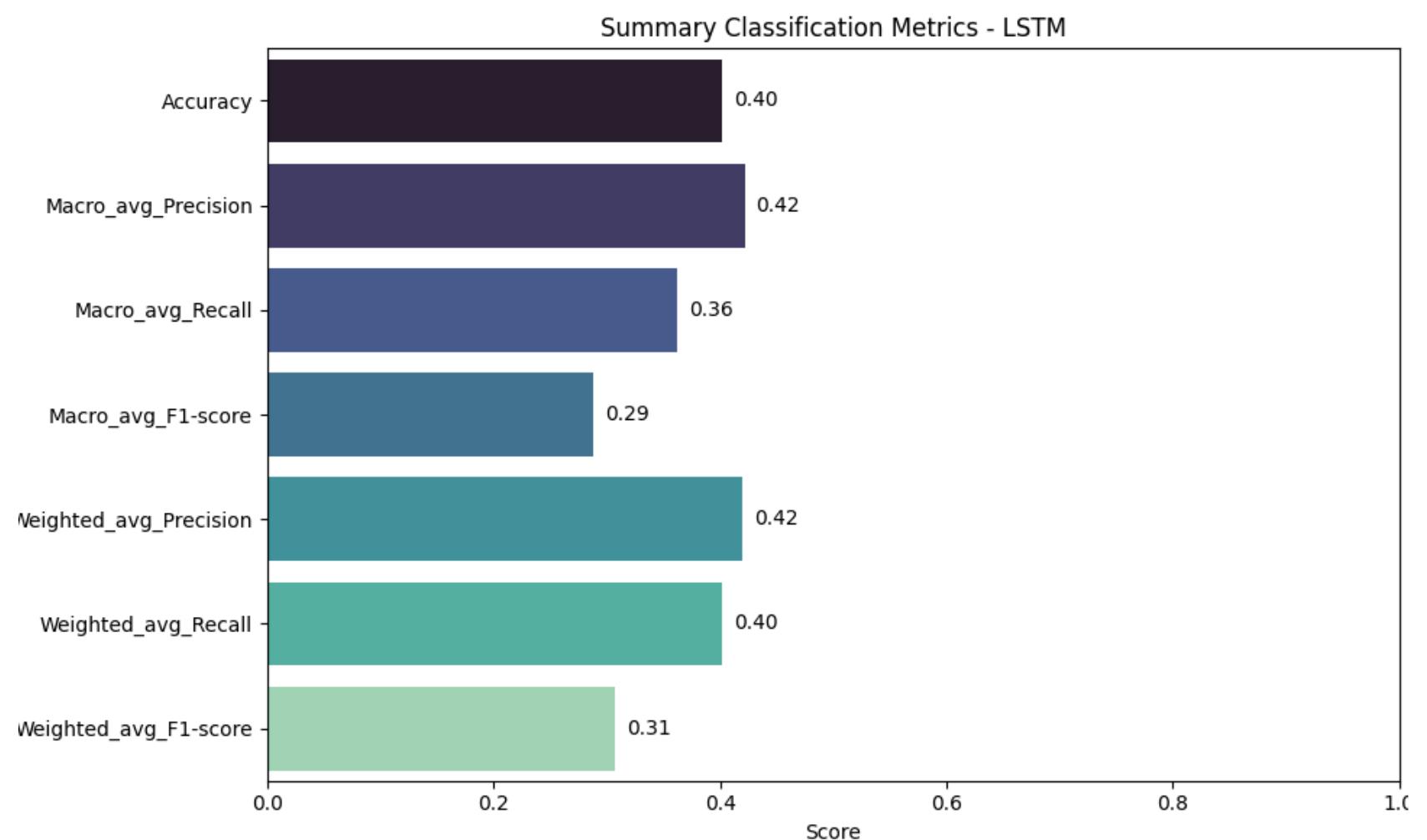


# Error Analysis





# Summary Metrics



# Insights & Discussion



**Temporal modelling alone (LSTM)  
(data imbalance)**

**Sentiment signal ( $\uparrow$  recall for HOLD)  
richer representation**

**Tree-based models performance  
Representative data**

**Overall performance  
Needs more experimentation**



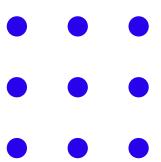
# Limitations & Next Steps

- Class imbalance (Focal loss, SMOTE-NC, cost-sensitive training)
- Limited features (macro-economic calendar, volatility indices, sentiment embeddings)
- Modeling: CNN-LSTM hybrids (local patterns)
- Robust evaluation with other real examples data



# Key Takeaways

- Pipeline reproducible end-to-end (data → features → models → reports).
- Sentiment adds context but might confuse
- XGBoost offers strong baseline, LSTM requires tuning & more data.
- Clear roadmap exists to reach >55 % macro-average score



# THANK YOU!